

INFORMACINĖS TECHNOLOGIJOS IR KALBA

Klaviatūros įtaka rašybos klaidoms, susijusioms su rašto ženklų vartojimu

Gintautas Grigas

Matematikos ir informatikos instituto docentas,
daktaras
Institute of Mathematics and Informatics,
Assoc. Prof., PHD.
Akademijos g. 4, LT-08663 Vilnius
Tel. (8 5) 210 93 44
El. paštas grigas@kti.mii.lt

Sigita Pedzevičienė

Matematikos ir informatikos instituto inžinierė
programuotoja
Institute of Mathematics and Informatics,
Engineer programmer
Akademijos g. 4, LT-08663 Vilnius
Tel. +370 684 44 008
El. paštas sigitab@kti.mii.lt

Analizuojama daromų rašybos klaidų, susijusių su vengimu vartoti raides su diakritiniais ženklais, priklausomybė nuo ženklų išdėstymo klaviatūroje. Analizei pasirinkta miestų rašyba pokalbių sistemos „Skype“ abonentų registracijos duomenyse. Palyginamos įvairių kalbų kompiuterio klaviatūros ir daromų klaidų skaičius, taip pat klaidos, susijusios su atskiromis tos pačios kalbos raidėmis, kurių rinkimo klaviatūra sudėtingumas skiriasi. Gauti rezultatai patvirtina hipotezę, kad daromų klaidų skaičius priklauso nuo raidžių rinkimo klaviatūra sudėtingumo.

Analizuojant pokalbių programų registracijos duomenis pastebėta, kad lietuviai daro daugiau rašybos klaidų negu kitų Europos tautų atstovai, vietoj raidžių su diakritiniais ženklais rašydami į jas panašias raides be diakritinių ženklų (Grigas, 2007). Buvo iškelta hipotezė, kad klaidų priežastis yra kompiuterio klaviatūra, nes Lietuvoje daugiausia naudojamose kompiuterio klaviatūrose raides su diakritiniais ženklais sunkiau surinkti, negu jų neturinčias. Tačiau to tyrimo tikslas buvo tik palyginti įvairių tautų atstovų daromų klaidų skaičių ir atkreipti lietuvių dėmesį, kad jie klaidų daro daugiau negu bet kuri kita tauta, ir paskatinti, kad darytų mažiau. Taip pat nebuvo analizuojamos kitų tautų klaviatūros.

Šio tyrimo tikslas – patikrinti iškeltą hipotezę atliekant išsamesnę klaidų analizę ir įvertinant klaviatūrų savitumus.

Analogiškų rašybos klaidų daro ir kitų tautų atstovai, tiksliai vieni daugiau, kiti mažiau.

Klaviatūros taip pat ne visų vienodos: vienose visos abėcėlės raidės laikomos lygiavertėmis, kitose raidės su diakritiniais ženklais laikomos antraeilėmis ir yra įdėtos į mažiau patogias pozicijas, dėl to jas rinkti darosi sudėtingiau. Todėl buvo analizuojamos įvairių kalbų klaviatūros pagal savitųjų tų kalbų raidžių rinkimą.

Buvo analizuojami pokalbių programos „Skype“ registracijos duomenys naudojantis metodika, aprašyta ankstesniame straipsnyje (Grigas, 2007). Siekiant vienareikšmiškumo apsiribota miestų vardų rašybos tikrinimu, nes kiekvienoje valstybėje jos miestų rašyba yra vienareikšmiška ir beveik kiekvienas sistemos abonentas jį nurodo (to negalima pasakyti apie vardus ir pavardes, kurių daugelis nenurodo arba pakeičia išgalvotais slapyvardžiais, o apie jų rašybos taisyklumą vienareikšmiškai negalima pasakyti). Siekiant didesnio išsamumo buvo tikrinama ir analizuojama tos pačios valstybės miestų rašyba, įėjus tų miestų

varduose naudojamos savitosios tos kalbos raidės, kurių rinkimo sudėtingumas aiškiai skiriasi. Taip pat buvo tikrinamos klaidos, daromos rašant miestų vardus iš mažųjų raidžių, kurių rinkimo sąlygos visų kalbų klaviatūrose yra vienodos. Šį tikrinimą galima pavadinti kontroliniu, nustatančiu kitų veiksmų įtaką daromų klaidų skaičiui.

Dėl ko lietuviai vengia rašyti raides su diakritiniais ženklais?

Kalbant su interneto naudotojais galima išgirsti nusiskundimų, kad raides su diakritiniais ženklais surinkti sunkiau arba jų turimame kompiuteryje iš viso neįmanoma. Kiti yra susitaikę su tokiu kompiuteriu ir manydami, kad be diakritinių ženklų galima apseiti, juos ignoruoja. Siekdami išvengti subjektyvumo, surengėme dvi nedideles apklausas. Vieną interneto svetainėje „Lietuvių kalba informacinėse technologijose“ (www.likit.lt), kitą – žmonių intensyviau naudojančių internetą, grupėse: vidurinės mokyklos 9, 11 ir 12 klasių mokinių, kolegijos ir universiteto studentų, informatikų. Paklausėme, kodėl elektroniniuose laiškuose ir internete jie nerašo raidžių su diakritiniais ženklais. Apklausos rezultatai pateikiami 1 lentelėje:

Labiau skiriasi tik 1 ir 4 atsakymų į klausimą rezultatas, matyt, dėl to, kad atsakymas „Rašau taisyklingsiai“ galėjo būti supras tas nevienareikšmiškai: ar visada rašau taisyklingsiai, ar tik kada nors rašiau netaisyklingsiai, ir tada žymiu kitą atsakymą. Tačiau pirmų trijų atsakymų į klausimą

santykis rodo, kad daugiau nei pusė apklaustųjų rašo netaisyklingsiai, o moksleivių ir studentų, rašančių netaisyklingsiai, yra dar daugiau.

Šia nedidele apklausa siekta ne nustatyti tikslą ir statistiškai pagrįstą situaciją Lietuvoje, o tik įsitikinti, kad tikslinga analizuoti raidžių rinkimo klaviatūra nepatogumus. Tai toliau ir darysime.

Kalbos ir klaviatūros

Jeigu tam tikrai raidei surinkti reikia daugiau pastangų, negu kitoms raidėms, tai sakome, kad ją surinkti sunkiau, ir tada kyla pagunda išvis tos raidės nevartoti. Sunkiau surinkti būna dėl to, kad: 1) raidės klavišas yra sunkiau pasiekiamoje viršutinėje skaitmenims skirtoje eilėje; 2) raidei surinkti reikia daugiau negu vieno klavišo paspaudimo (raidė yra antrame arba trečiame klaviatūros lygyje, raidė renkama su kombinaciniu klavišu ir pan.); 3) ant klavišo nėra užrašytos raidės, nors jos vieta intuityviai aiški; 4) ant klavišo nėra užrašytos raidės ir jos vieta intuityviai neaiški; 5) užrašo vieta ant klavišo ne vietoje – neatitinka ženklo lygio; 6) klaviatūra neturi kitų dialogui su kompiuteriu reikalingų ir privalomų pagal ISO standartą (ISO 9995, 2006) ženklų ir dėl to vengiama ją naudoti (tai netiesioginis nepatogumas ir klaviatūros nevisavertiškumo požymis).

Šie požymiai įvairiai susipina skirtingų kalbų klaviatūrose ir juos kiekybiškai įvertinti sunku. Todėl klaviatūras skirstysime į didesnes, aiškiai atskiriamas tris grupes:

A. Visavertė originali klaviatūra, kurioje visų

tos kalbos savitųjų raidžių rinkimas neturi nė vieno čia išvardyto rinkimą sunkinančio požymio. Šiai grupei priklauso šių toliau nagrinėjamų kalbų klaviatūros: danų, estų, suomių, švedų, vokiečių.

B. Iš dalies visavertė originali klaviatūra, kurioje dalis savitųjų raidžių neturi rinkimą sunkinančių požymių, o dalis – turi. Dažniausiai tokią klaviatūrą naudoja

1 lentelė. Apklausos apie rašybos klaidų priežastis rezultatai

Eil. Nr.	Atsakymas	Likit.lt		Grupės		Bendras	
		Žmonių skaičius	%	Žmonių skaičius	%	Žmonių skaičius	%
1	Raides su diakritiniais ženklais sunkiau surinkti	89	40,1	76	67,9	165	49,4
2	Mano kompiuteryje neįmanoma surinkti raidžių su diakritiniais ženklais	7	3,2	4	3,6	11	3,3
3	Manau, kad raidės su diakritiniais ženklais lietuvių kalboje nereikalingos	14	6,3	4	3,6	18	5,4
4	Rašau taisyklingsiai	112	50,5	28	25,0	140	41,9
	Iš viso	222	100	112	100	334	100

kalbos, kurių abėcėlės turi daug raidžių (pvz., čekų, turinti 42 raides). Šiai grupei priklauso čekų, islandų, ispanų, lenkų, prancūzų klaviatūros.

C. Pakoreguota kitos kalbos klaviatūra, kurioje kiekviena savitoji raidė turi vieną ar daugiau čia išvardytų rinkimą sunkinančių požymių. Šiai grupei priklauso latvių ir lietuvių klaviatūros. Abi kalbos naudoja amerikietišką fizinę klaviatūrą.

Antroje lentelėje pateikti rašybos klaidų, daromų pokalbių programos „Skype“ registracijos duomenyse užrašant savų miestų vardus, tikrinimo rezultatai, išreikšti taisyklingai užrašytų vardų skaičiais ir procentais. Miestų vardai pateikiami originalo kalbomis. Duomenys pavaizduoti grafiškai 1 paveiksle.

2 lentelė. Miestų vardų rašyba

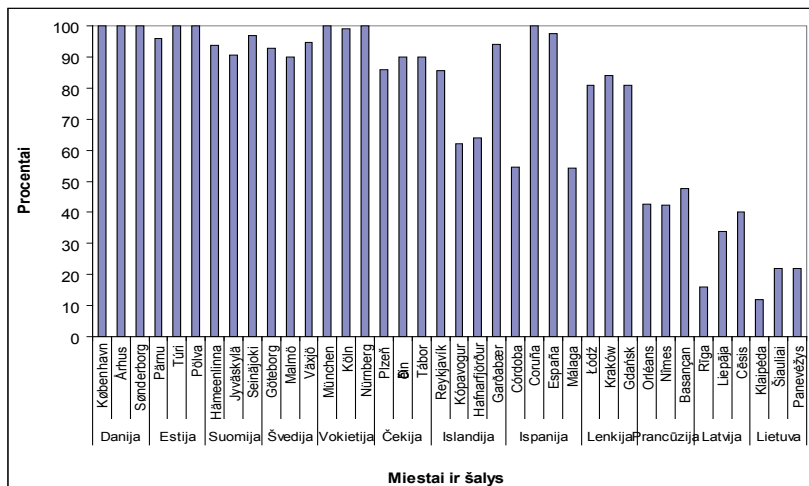
Klaviatūros grupė	Kalba	Miestas	Iš viso	Taisyklingų %	Taisyklingų %	Bendras %
A	Danų	København	512	512	100	100
		Århus	198	198	100	
		Sønderborg	37	37	100	
A	Estų	Pärnu	146	140	96	97
		Türi	17	17	100	
		Põlva	15	15	100	
A	Suomių	Hämeenlinna	80	75	94	93
		Jyväskylä	181	164	91	
		Seinäjoki	65	63	97	
A	Švedų	Göteborg	459	426	93	92
		Malmö	333	300	90	
		Växjö	165	156	95	
A	Vokiečių	München	682	679	100	100
		Köln	397	395	99	
		Nürnberg	179	179	100	
B	Čekų	Plzeň	494	423	86	87
		Děčín	155	140	90	
		Tábor	139	125	90	
B	Islandų	Reykjavík	342	293	86	79
		Kópavogur	117	73	62	
		Hafnarfjörður	80	51	64	
B	Ispanų	Garðabær	67	63	94	62
		Córdoba	123	67	54	
		Coruña	68	68	100	
B	Lenkų	España	40	39	98	82
		Málaga	401	217	54	
		Łódź	555	448	81	
B	Prancūzų	Kraków	514	430	84	45
		Gdańsk	416	335	81	
		Orléans	82	35	43	
C	Latvių	Nîmes	104	44	42	23
		Basançan	267	127	48	
		Rīga	862	137	16	
C	Lietuvių	Liepāja	304	103	34	17
		Cēsis	134	54	40	
		Klaipėda	649	75	12	
C	Lietuvių	Šiauliai	420	92	22	17
		Panevėžys	303	66	22	

Į sąrašą įtrauktos visos kalbos, analizuotos straipsnyje (Grigas, 2007), taip pat panaudoti minėto straipsnio duomenys. Sąrašas papildytas keturiomis naujomis kalbomis. Suomių ir švedų kalbos įtrauktos dėl to, kad ankstesniame darbe nebuvo nė vienos skandinavų kalbos, ispanų kalba – dėl to, kad Ispanijoje yra miestų vardų, turinčių raidžių su skirtingais rinkimą sunkinančiais požymiais, prancūzų kalba – dėl to, kad ji yra viena labiausiai paplitusių Europoje.

Siekiant išvengti priklausomybės nuo kalbos mokėjimo ir nuo techninių priežasčių, galinčių atsirasti svetur gyvenančiam žmogui, buvo registruojami duomenys tik tų žmonių, kurie savo duomenyse nurodė, kad susirašinėja analizuojama kalba ir gyvena valstybėje, kurioje ta kalba yra pagrindinė.

Kaip matome iš lentelės, daromų klaidų skaičius visiškai atitinka kalbų grupes, suskirstytas pagal rinkimą sunkinančius požymius.

Grupių viduje kalbos išdėstytos pagal abėcėlę. Į C grupę patenka latvių ir lietuvių kalbų klaviatūros. Latvijoje naudojama amerikietiška klaviatūra papildyta savitosiomis latvių kalbos raidėmis, kurių yra dvylika: ĀČĒĢĪĶĻŅŠŪŽ. Visos jos išdėstytos trečiame (mažosios) ir ketvirtame (didžiosios) klaviatūros lygiuose. Ant klavišų nėra užrašytos, bet jų vieta suvokiama intuityviai: raidės su diakritiniu ženklu klavišas yra tas pats, kaip ir ją atitinkančios raidės be diakritinio ženklo (pvz., raidė Ū ant klavišo su užrašu U). Atitikimas abipusiai viena-reikšmis, nes kiekviena raidė su diakritiniu ženklu turi tik vieną ją atitinkančią raidę be diakritinio ženklo.



1 pav. Taisyklingai, vartojant raides su diakritiniais ženklais, užrašytų miestų vardų procentai

Lietuvoje naudojama amerikietiška klaviatūra, papildyta savitosiomis lietuvių kalbos abėcėlės raidėmis, kurių yra devynios: AČĖĖIŠŲŪŽ. Visos jos išdėstytos viršutinėje klavišų eilėje vietoj skaitmenų ir kitų šioje JAV klaviatūroje esančių ženklų. Yra tvarkyklių, kuriose prarasti ASCII ženklai atkurti trečiame ir ketvirtame klaviatūros lygiuose. Paplitusios dvejopos fizinės klaviatūros: su užrašytais raidėmis ant klavišų, bet netinkamoje vietoje – ten, kur turėtų būti rašomi trečio klaviatūros lygio ženklai, ir be užrašytų. Pastaruoju atveju raidės klavišas intuityviai nėra aiškus (pvz., raidė Ū yra ant klavišo su užrašu 8).

Čia aptarėme tik pagrindines ir visuotinai naudojamas klaviatūras, kurios operacinėse sistemose „Windows XP“ ir „Windows Vista“ laikomos numatytosiomis. Yra ir daugiau klaviatūrų (pvz., lietuviška standartinė klaviatūra, kuri patektų į pirmą mūsų aprašytą klaviatūrų grupę), bet jos naudojamos tiek mažai, kad neturi įtakos klaidų procentui.

Didžiųjų raidžių klaidos

Mūsų atlikta apklausa buvo lietuviškoje interneto svetainėje ir natūralu, kad jos rezultatai atspindi tik lietuvių požiūrį. Lyginant įvairias kalbas gali atsirasti kitokių skirtumų, priklausančių nuo konkrečios tautos: žmonių tvarkin-

gumo, pagarbos savo kalbai ir pan. Todėl buvo atliktas kontrolinis eksperimentas su raidėmis, kurių rinkimo sąlygos vienodos visų valstybių klaviatūrose. Tada ir daromų klaidų skaičius neturėtų priklausyti nuo klaviatūros.

Būdinga klaida yra miesto vardo rašymas iš mažosios raidės. Norint surinkti didžiąją raidę, reikia papildomai paspausti dar

vieną – antro lygio – klavišą. Tai paprasčiau, negu surinkti pasinaudojant trečio lygio klavišu, nes antro lygio klavišai šiek tiek didesni, jie yra du. Papildomas darbas nedidelis, bet kai kas jį pritingi atlikti ir savo miesto vardą parašo iš mažosios raidės. Čia vienodos ir visų kalbų gramatikos taisyklės, reikalaujančios miestų vardus rašyti iš didžiųjų raidžių.

Eksperimentui pasirinkome valstybių sostinių vardus – bent iš pagarbos sostinei jos vardą derėtų rašyti iš didžiosios raidės. Rezultatai pavaizduoti 2 paveiksle.

Kaip matome, šios rūšies klaidų daro visi, o skirtumai tarp tautų nedideli. Rezultatų išsibarsčymas, išreikštas vidutiniu kvadratinu nuokrypiu, yra 5,5, o klaidų dėl diakritinių ženklų išsibarsčymas, apskaičiuotas iš 2 lentelės, yra 29,6. Tai dar vienas hipotezės, jog klaidos priklauso nuo raidžių rinkimo sudėtingumo, patvirtinimas.

Klaidos su tos pačios kalbos raidėmis, kurių rinkimo sudėtingumas skirtingas

Yra klaviatūrų (Keyboard, 2009; Windows, 2009), kuriose tos pačios kalbos savitosios raidės rinkimo sunkumo požiūriu išdėstytos skirtingose pozicijose. Pavyzdžiui, islandų raidė Ð yra viršutinėje raidžių eilėje greta raidės P. Todėl ją rinkti nėra sudėtingiau negu bet kurią kitą. Raidė Í renkama paėiliui paspaudžiant du

klavišus: dešiniojo kirčio ženklą (akūto) ir raidės I klavišus. Taigi ją rinkti sudėtingiau. Analogiškų skirtumų yra ir kai kurių Ispanijos miestų varduose.

Pasinaudodami 2 lentelės duomenimis grafiškai pavaizduosime Islandijos ir Ispanijos miestų varduose daromas klaidas, kiekvienos valstybės miestus suskirstę į dvi grupes pagal tai, ar savitosios raidės rinkimo sudėtingumas yra toks pats, kaip ir raidžių su diakritiniais ženklais, ar didesnis (3 pav.).

Kaip matome, rezultatai dar kartą patvirtina, kad kuo sudėtingesnis raidžių rinkimas, tuo dažniau daroma klaidų.

Išvada

Klaidų darymas rašant raides su diakritiniais ženklais priklauso nuo klaviatūros: kuo sudėtingiau raidę surinkti, tuo dažniau vietoj jos rašoma raidė be diakritinio ženklo. Tai būdinga įvairioms kalboms, vartojančioms lotyniškus rašmenis ir turinčioms raidžių, nesančių pagrindinėje lotynų abėcėlės dalyje (anglų kalbos abėcėlėje).

LITERATŪRA

GRIGAS, G.; PEDZEVIČIENĖ, S. (2007) Asmenvardžių ir vietovardžių rašybos klaidos pokalbių programų registracijos duomenyse. *Informacijos mokslai*, t. 42–43, p. 141–144.

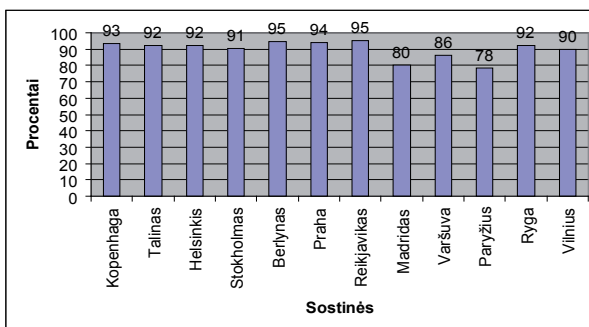
ISO/IEC 9995 (2006). Information technology – keyboard layouts for text and office systems. Part 1–8. ISO.

THE RELATIONS BETWEEN KEYBOARD LAYOUT AND SPELLING ERRORS OF PARTICULAR LANGUAGE SPECIFIC LETTERS

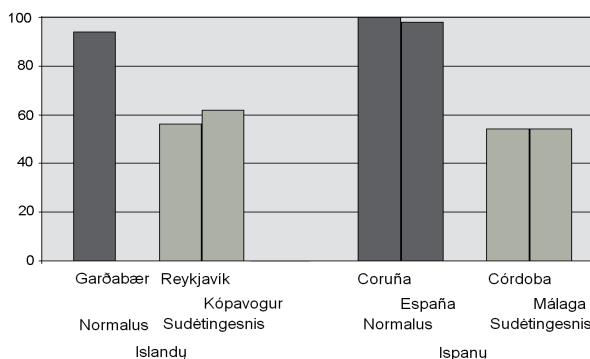
Gintautas Grigas, Sigita Pedzevičienė

Summary

The relation between the keyboard layout of language specific letters, usually accented letters, and spelling error frequency is investigated. Names of cities in login data of instant messenger Skype are chosen for investigation. Keyboards of 12 European languages using Latin script are related to the num-



2 pav. Miestų vardų, parašytų iš didžiosios raidės, procentai



3 pav. Islandijos ir Ispanijos miestų vardų rašyba, kai raidžių rinkimo klaviatūra sudėtingumas yra skirtingas

KEYBOARD layout (2009). Wikipedia [žiūrėta 2009 m. gegužės 21 d.]. Prieiga per internetą: http://en.wikipedia.org/wiki/Keyboard_layout.

WINDOWS keyboard layouts (2009) [žiūrėta 2009 m. gegužės 21 d.]. Prieiga per internetą: [http://msdn.microsoft.com/lt-lt/goglobal/bb964651\(en-us\).aspx](http://msdn.microsoft.com/lt-lt/goglobal/bb964651(en-us).aspx).

ber of spelling errors. Similar comparisons are made for letters requiring different efforts for typing within the same language. The obtained results confirm hypothesis that the number of deliberately made errors depends on the complexity of typing letters.